

The renormalisation group via statistical inference

Cédric Bény¹ and Tobias J. Osborne¹

¹*Institut für Theoretische Physik, Leibniz Universität Hannover, Appelstraße 2, 30167 Hannover, Germany*

(Dated: February 21, 2014)

In physics one attempts to infer the rules governing a system given only the results of imperfect measurements. Hence, microscopic theories may be effectively indistinguishable experimentally. We develop an operationally motivated procedure to identify the corresponding equivalence classes of theories. Here it is argued that the *renormalisation group* arises from the inherent ambiguities in constructing the classes: one encounters *flow parameters* as, e.g., a regulator, a scale, or a measure of precision, which specify representatives of the equivalence classes. This provides a unifying framework and identifies the role played by information in renormalisation. We validate this idea by showing that it justifies the use of low-momenta n -point functions as relevant observables around a gaussian hypothesis. Our methods also provide a way to extend renormalisation techniques to effective models which are not based on the usual quantum-field formalism, and elucidates the distinctions between various type of RG.

The renormalisation group (RG), as conceived by Wilson [1, 2], relies on the idea that it is possible to describe long-distance physics while essentially ignoring short-distance phenomena; Wilson argued that, if we are content with predictions to some specified accuracy, the effects of physics at smaller lengthscales can be absorbed into the values of a few parameters of some *effective theory* for the long-distance degrees of freedom. The RG now underpins much of our understanding of modern theoretical physics and provides the interpretational framework for quantum field theories. It has been applied in a dazzling array of incarnations to study systems from statistical physics [3] to applied mathematics [4].

The general applicability of RG techniques strongly suggests the existence of a deep unifying principle which would make it possible to directly compare different manifestations of the RG and to unlock its full potential. It has been suggested that such a general implementation-independent formulation of the RG is to be found in an information-theoretic approach [5] because the RG works by *ignoring* certain aspects of the system. Although the information-theoretic flavour of the RG is manifest in the case of block-decimation [6–8], it is far less obvious in the context of particle physics from where the terminology of renormalisation originates [9]. Previous attempts at tackling this problem (see, e.g., [10–14] for a selection) depended heavily on details of the chosen model or formalism and do not offer the truly general unification that one might hope for.

The objective of this paper is to propose an operationally motivated, model-independent, and information-theoretic framework for the RG. Our main result is a precise description of the renormalisation group, as implemented in terms of a regulator (as found in QFT), in terms of information-theoretic quantities.

In pursuing this goal we found it necessary to first step back and reconsider the task of *inference* in quantum mechanics. We begin by phrasing this task as a game played between two players: a passive one, Alice, who possesses a quantum system, and Bob, who perceives the system via a known noisy quantum channel \mathcal{E} . (A channel is a

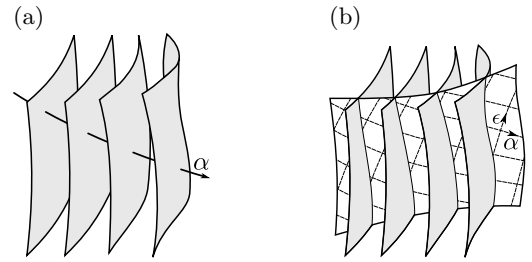


FIG. 1. The shaded planes represent equivalence classes of states which cannot be distinguished experimentally. They are intersected by the manifold of effective theories, parametrised in example (a) by a sole parameter α , and, in example (b), additionally by a regularisation parameter ϵ . The intersection lines are renormalisation trajectories $\alpha(\epsilon)$.

linear map on density matrices which preserves positivity and normalisation, even as part of a larger system.) Alice’s state represents the true state of a physical system, while Bob is an experimentalist whose experimental limitations are formalised by the channel. When Bob tries to infer the state of Alice’s system, he is faced with the ill-posed *inverse problem* of inverting a quantum channel to find the input from the output. This task is not well-posed because distinct states of Alice may yield states which are nearly indistinguishable for Bob.

Let us consider first a situation where the channel has a non-trivial kernel. For instance, \mathcal{E} could be the partial trace over all high-momentum modes of a theory. If two states ρ and ρ' are such that $\mathcal{E}(\rho) = \mathcal{E}(\rho')$ then they cannot be distinguished by Bob and hence are both just as good as hypotheses for Alice’s state. This indistinguishability results in *equivalence classes* of states: all that Bob can hope to do is to determine in which equivalence class the true state is. The classes can be parameterised by a smooth manifold of unique representatives (Fig. 1a). For instance, if \mathcal{E} traces out high-momentum modes, the equivalence classes can be labelled by states whose high momentum modes are in some fiducial product state.

Once the equivalence classes of experimentally indis-

tinguishable states are identified, we propose that the various renormalisation groups may all be interpreted as resulting from the exploration of the freedom available in choosing the representative within a class. For example, by modifying a regularisation parameter, as occurs in high-energy physics, or by simplifying the description of the state and isolating the relevant degrees of freedom, as commonly practiced in condensed matter theory. Before we describe these two cases in more depth, we want to consider more general, and more realistic experimental limitations involving, for instance, both a limit on spatial resolution and on the ability to resolve local degrees of freedom. This requires us to consider states which are only *approximately* indistinguishable by Bob.

General framework — A reasonable measure of distinguishability between two states ρ and ρ' to be used in this situation is the relative entropy $S(\rho' \parallel \rho)$, which measures the optimal exponential rate of decrease of the probability of mistaking ρ' for ρ as a function of the number of copies available, while still letting the probability of mistaking ρ for ρ' go to zero [15]. This interpretation is for an observer able to measure any observable on Alice's system. Bob, however, can only measure observables of the form $\mathcal{E}^\dagger(A)$, so the relevant measure for him is instead $S(\mathcal{E}(\rho') \parallel \mathcal{E}(\rho))$.

Consequently, we could attempt to deem two states ρ and ρ' experimentally equivalent if $S(\mathcal{E}(\rho') \parallel \mathcal{E}(\rho)) < \delta$ for some desired maximal rate δ . However, this does not define an equivalence relation. Nevertheless, if δ is sufficiently small, we still expect the set of states ρ' close to ρ in this sense to form an approximately linear subspace of matrices, as occurs in the case $\epsilon = 0$. This motivates us to *linearise* the relation around a starting hypothesis ρ .

Let's write $\rho' = \rho + \epsilon X$, where we call the tangent operator X , which must be hermitian and traceless, a *feature*. Then, to lowest order in ϵ , we have

$$S(\rho + \epsilon X \parallel \rho) \approx \epsilon^2 \text{Tr}(X \Omega_\rho^{-1}(X)) \equiv \epsilon^2 \langle X, X \rangle_\rho,$$

where $\Omega_\rho^{-1}(Y) = \frac{d}{dt} \log(\rho + tY)|_{t=0}$ is a non-commutative version of the operation “division by ρ ”. The quantity $\langle X, Y \rangle_\rho \equiv \text{Tr}(X \Omega_\rho^{-1}(Y))$ is an inner product on operators and allows us to measure not only the “length” or “size” of a feature, but also the “angle” between two features X and Y — it is a *metric* in the sense of differential geometry and is one of the many quantum generalisations of the Fisher information metric [16].

In this linear approximation, a state $\rho + X$ is approximately indistinguishable from $\rho + Y$ by Bob if $\langle \mathcal{E}(X - Y), \mathcal{E}(X - Y) \rangle_{\mathcal{E}(\rho)} < \delta$. Although this relation between X and Y is symmetric, it is still not an equivalence relation because it is not transitive. The set of states $\rho + X$ satisfying this condition is an ellipsoid within Alice's state space. If \mathcal{E} has a non-trivial kernel, for instance, this ellipsoid is infinitely wide in the null directions. Consequently, in the generic case we use the following idealised relation: we say that the two states $\rho + X$ and $\rho + Y$ are equivalent if $X - Y$ lies in the span of the

“largest” principal directions of the ellipsoid (called *irrelevant*), which is equivalent to saying that $X - Y$ is orthogonal to the span of the n smallest principal directions of the ellipsoid (called the relevant directions).

This idealisation removes any trace of the desired precision δ , as we are only talking of the direction of $X - Y$ independently of its magnitude. Instead, Bob must choose the number n of smallest principal directions he deems relevant. A pertinent way of doing this is to consider the case where the channel \mathcal{E} depends on a parameter σ parameterising the precision of Bob's instruments, and to worry about the asymptotic behaviour of the norm $\langle \mathcal{E}(Z), \mathcal{E}(Z) \rangle_{\mathcal{E}(\rho)}$ in the limit of large imprecision σ . The choice of the relevance threshold n then amounts to choosing the type of asymptotic behaviour that we deem negligible. Note: as long as σ is in fact finite, one must remember that no matter what n is, an irrelevant feature $Z = X - Y$ would become observable if it were to have norm $\langle \mathcal{E}(Z), \mathcal{E}(Z) \rangle_{\mathcal{E}(\rho)} > \delta$, in which case our idealisation fails.

The principal directions of the ellipsoid can be found by solving an eigenvalue equation. Let \mathcal{R}_ρ be the super-operator defined by $\langle X, \mathcal{R}_\rho(Y) \rangle_\rho = \langle \mathcal{E}(X), Y \rangle_{\mathcal{E}(\rho)}$ for all features X, Y [17]. Explicitly, $\mathcal{R}_\rho = \Omega_\rho \mathcal{E}^\dagger \Omega_{\mathcal{E}(\rho)}^{-1}$. Classically, \mathcal{R}_ρ implements *Bayesian inference* with prior ρ ; we are comparing Alice's feature X with that Bob can infer from his data. The *principal features* X_j are the solution of $\mathcal{R}_\rho \mathcal{E}(X_j) = \eta_j X_j$ with eigenvalues $1 \geq \eta_1 \geq \eta_2 \geq \dots \geq 0$. We call η_j the *relevance* of X_j [18]. The linear operator $\mathcal{R}_\rho \mathcal{E}$ is self adjoint in the scalar product $\langle \cdot, \cdot \rangle_\rho$, as well as positive. Therefore, the principal features X_j form a complete orthogonal basis of the tangent space at ρ .

We call a feature X *relevant* if it is in the span of the n most relevant principal features X_1 to X_n , and *irrelevant* if it is orthogonal to those, i.e., in the span of the remaining principal features. Our idealised equivalence classes consider $\rho + X$ and $\rho + Y$ to be equivalent from the point of view of Bob if and only if $X - Y$ is irrelevant, or, equivalently, if $\langle X - Y, Z \rangle_\rho = 0$ for all relevant feature Z .

In order to obtain a physically more intuitive condition, let us define the *principal observables* to be the operators $A_j = \Omega_\rho^{-1}(X_j)$. Analogously, we say that A is relevant if it is in the span of the n most relevant principal observables (and therefore of the form $A = \Omega_\rho^{-1}(X)$ where X is relevant). The principal observables A_j are the solutions of the dual Heisenberg-picture eigenvalue equation

$$\mathcal{E}^\dagger \mathcal{R}_\rho^\dagger(A_j) = \eta_j A_j. \quad (1)$$

With this definition, our equivalence condition amounts to considering two effective states ρ' and ρ'' to be equivalent (in the neighbourhood of ρ) when they yield the same expectation values for *all relevant observables*:

$$\rho' \sim_\rho \rho'' \quad \text{iff} \quad \text{Tr}(\rho' A_j) = \text{Tr}(\rho'' A_j) \quad \text{for } j \leq n.$$

For instance, consider the strictest possible relevance threshold where only features with exactly zero relevance

are deemed to be irrelevant. These are the operators X in the kernel of \mathcal{E} . In this case we recover the exact state-independent equivalence relation which identifies $\rho' \sim \rho''$ if $\mathcal{E}(\rho') = \mathcal{E}(\rho'')$. The corresponding relevant observables are the self-adjoint operators A satisfying $\text{Tr}(AX) = 0$ for all X in the kernel of \mathcal{E} , which are precisely those of the form $\mathcal{E}^\dagger(B)$ for some B . Observe that, given Eq. 1, all observables with non-vanishing relevance are indeed of this form. In addition, these are all the observables that Bob can ever hope to measure expectation values of, since for all B , $\text{Tr}(B\mathcal{E}(\rho)) = \text{Tr}(\mathcal{E}^\dagger(B)\rho)$.

One classical mode — For a simple but nontrivial example suppose that Alice has a stochastic classical system consisting of a single real variable, e.g., the position x of a particle. The true state to be discovered by Bob is a probability distribution on \mathbb{R} : $x \mapsto \rho(x)$. Bob’s experimental limitation consists of a finite precision σ at which he can resolve the particle’s position. This can be modelled by a channel \mathcal{E} —in this case a stochastic map since the system is classical—whose effect is a convolution of Alice’s probability distribution with a Gaussian of width σ . Suppose, further, that Bob’s initial hypothesis is a simple gaussian distribution, which we think of as a thermal state $\rho(x) \propto e^{-\beta H(x)}$ for the “hamiltonian” $\beta H(x) = \frac{x^2}{2\tau^2}$. Our eigenvalue equation can be solved for this system in the heisenberg picture, yielding the hermite polynomials $\text{He}_n(x/\tau)$ with eigenvalues $\eta_n = [\tau^2/(\sigma^2 + \tau^2)]^n$, or $\eta_n \approx (\tau/\sigma)^{2n}$ for $\sigma \gg \tau$. Since the first n hermite polynomials span all degree- n polynomials this means that two nearby states are equivalent from the point of view of Bob exactly when they have the same first n moments, where n is the threshold chosen by Bob.

Following our framework, suppose that Bob chooses the threshold $n = 2$, which means that he deems two states to be equivalent if their distinguishability vanishes at order $\sigma^{2(n+1)} = \sigma^6$. Then the hamiltonian $\beta H_0(x) = \frac{x^2}{2\tau_0^2} + \lambda x^4$ is equivalent to $\beta H_1(x) = \frac{x^2}{2\tau_1^2}$ provided that τ_1 is “renormalised” so that the two corresponding states have the same second moment. We see that, as the threshold n decreases (and hence more information is neglected) an initially complex model can be simplified. This corresponds to the type of renormalisation group employed in condensed matter theory. It is important to note that, although this example of simplification is quite natural—we set the redundant parameter λ to zero, hence removing a term in the hamiltonian—it is in many ways arbitrary, as any other state with the same first two moment would be as good an effective state.

The situation in particle physics is quite different. Quantum field theories typically come with an unwanted parameter, a *regulator* ϵ , which has no true physical significance except that it is not possible to remove it mathematically. At best it may mimic a lattice spacing, needed in order to avoid high energy modes from making some observable diverge. But it could also be a much less physical fractional dimension of space used as a convenient

way to make spherical integrals converge.

The presence of this unphysical parameter, however, is not a problem if the observable predictions of the theory do not depend on it. This is possible if we assume some reasonable limitation on Bob’s measurement abilities, so that any change in ϵ can be compensated by a change in the state’s other parameters so as to stay within a given equivalence class (Fig. 1b). This dependance of the state’s parameters on ϵ is the type of RG flow which naturally occurs in quantum field theory.

We can again use our simple classical example in order to illustrate this phenomenon. Suppose that Bob works with the relevance threshold $n = 4$, and treats the parameter λ perturbatively to first order: $\rho'(x) = e^{-x^2/\tau^2 - \lambda x^4} \approx e^{-x^2/\tau^2} (1 - \lambda x^4)$. Using this perturbative approach he may well measure λ and find that a small *negative* value fits his data nicely. However, if Bob were to then believe this characterises the true state of Alice’s system via a hamiltonian $H'(x) = x^2/\tau^2 + \lambda x^4$ he is in for some trouble because the corresponding thermal state cannot be normalised. This is different from the type of divergence occurring in QFT, but it is good enough for our analogy. Since any state which shares the same first four moments would be indistinguishable for Bob, he has a lot of freedom to fix his theory. For example, he can add a *regularisation* term of the form ϵx^6 to the hamiltonian, which makes the state well defined no matter how small ϵ is. The effect of this term on the second and fourth moment of the state can be compensated by appropriately modifying the parameters τ and λ to $\tau(\epsilon)$ and $\lambda(\epsilon)$.

This flow of the effective state as a function of a regularisation parameter ϵ , which is the logical equivalent of the RG flow in QFT, has no *a priori* relationship to the flow generated by varying the threshold n . However, in quantum field theory, divergences can be identified as contributions from an infinite number of irrelevant features. Therefore, if the theory is regularised by subtracting the irrelevant features from the state, then the two types of RG flow can coincide (see below).

One quantum mode — Before applying our formalism to a quantum field theory, we first examine a single quantum mode. We consider the quantum system with Hilbert space $L^2(\mathbb{R})$, in the neighbourhood of a gaussian state with characteristic function $\chi_\rho(x, p) = e^{-\frac{1}{4}(u^2 x^2 + v^2 p^2) + p_0 x - x_0 p}$. A lack of precision in measuring the position \hat{x} and momentum \hat{p} can be formalised as a gaussian channel which maps u^2 to $u^2 + \sigma_p^2$ and v^2 to $v^2 + \sigma_x^2$, where σ_x and σ_p are the uncertainties in measuring \hat{x} and \hat{p} respectively. Knowing that Ω_ρ is the operator derivative of the exponential function, and Ω_ρ^{-1} the derivative of the logarithm, it is easy to see that, to lowest order in ϵ , $\mathcal{E}(e^{-H+\epsilon A}) \simeq e^{-H'+\epsilon \mathcal{R}_\rho^\dagger(A)}$ where $\rho \propto e^{-H}$ and $\mathcal{E}(\rho) \propto e^{-H'}$. Since we know how the gaussian channel \mathcal{E} acts on gaussian states, this allows us to easily evaluate the effect of \mathcal{R}_ρ^\dagger on quadratic polynomials in \hat{x} and \hat{p} . We find that both \hat{x} and \hat{p} are principal observables (eigenvectors of $\mathcal{E}^\dagger \mathcal{R}_\rho^\dagger$). Asymptotically for large σ_x and

σ_p , their relevances are $\eta(\hat{x}) \approx \frac{v}{su} \sigma_x^{-2}$ and $\eta(\hat{p}) \approx \frac{u}{sv} \sigma_p^{-2}$, where we used $s = \coth^{-1}(uv)$. In terms of u and v , the second-order principal observables are more complicated linear combinations of \hat{x}^2 , \hat{p}^2 and $\mathbf{1}$, even asymptotically for large σ_x and σ_p . However, if the state ρ is very mixed ($uv \rightarrow \infty$), then we find the eigenvectors $\hat{x}^2 - \frac{s}{2} \frac{u}{v} \mathbf{1}$ and $\hat{p}^2 - \frac{s}{2} \frac{v}{u} \mathbf{1}$ with respective eigenvalues $u^4 \sigma_p^{-4}$ and $v^4 \sigma_x^{-4}$. We conjecture that all observables orthogonal to those have relevance of order at least 6 in $1/\sigma_p$ and $1/\sigma_x$.

Quantum fields — In general, a quantum Gaussian channel is parametrised by two real matrices X and Y , so that its effect on a Gaussian state's covariance matrix γ is $\gamma \mapsto X^T \gamma X + Y$. Correspondingly, the expected field ϕ_0 , if nonzero, is mapped to $X\phi_0$. In order to formalise Bob's finite spatial precision in the context of a field theory, one may use an operator X whose effect is to randomly swap spatial modes according to a gaussian probability distribution with variance σ . In addition, a lack of precision in measuring *local field values* can be simulated with a matrix Y which is proportional to the identity on the field coordinates and on the field canonical conjugates, but with different coefficients. Because X is translation invariant we see that in the neighbourhood of a translation-invariant quadratic theory the effect of this channel factors for each momentum mode. On each mode, the effect of the channel is similar to that analysed above.

For concreteness, we consider a scalar field theory. The hamiltonian defining ρ can be written as $H = \frac{1}{2} \int dk (\Pi_k^2 + \omega_k^2 \Phi_k^2)$, where $\omega_k = \sqrt{k^2 + m^2}$ and the operators Π_k and Φ_k are self-adjoint canonical conjugates. The effect of the gaussian channel \mathcal{E} on states of the form $\rho \propto e^{-H'}$ where $H' = \int dk \coth^{-1}(u_k v_k) \left(\frac{v_k}{u_k} (\Pi_k - \delta_k \mathbf{1})^2 + \frac{u_k}{v_k} (\Phi_k - \epsilon_k \mathbf{1})^2 \right)$, is to map u_k^2 to $X_k^2 u_k^2 + 2h_\Phi^2$ and v_k^2 to $X_k^2 v_k^2 + 2h_\Pi^2$, and δ_k to $X_k \delta_k$ and ϵ_k to $X_k \epsilon_k$, where $X_k = e^{-\frac{1}{2} k^2 \sigma^2}$, and h_Π and h_Φ parameterise the precision at which the fields are resolved. Using the same trick as for a single quantum mode, we obtain that, asymptotically for $h_\Phi h_\Pi \gg 1$, the field operators Φ_k and Π_k are principal observables with respective relevances $\eta(\Phi_k) \simeq 1/(\eta_0^{-1} + \beta \omega_k^2 h_\Phi^2 e^{k^2 \sigma^2})$ and $\eta(\Pi_k) \simeq 1/(\eta_0^{-1} + \beta h_\Pi^2 e^{k^2 \sigma^2})$, where $\eta_0 := \frac{\beta \omega_k}{2} \coth \frac{\beta \omega_k}{2}$.

Since the channel acts independently on each mode we find that the products $\Phi_{k_1} \dots \Phi_{k_n}$ have relevance $\eta(\Phi_{k_1}) \dots \eta(\Phi_{k_n})$, provided that the momenta k_1, \dots, k_n are all distinct. In order to define which observables are relevant, we have to choose the desired asymptotic behaviour of their relevance in terms of the three noise parameters σ , h_Π and h_Φ . Since the relevance always decays exponentially in $\sum_i k_i^2 \sigma^2$, the product of field operators are essentially irrelevant as soon as they involve operators with mode $k > 1/\sigma$. The behaviour of the relevance in term of h_Π and h_Φ is polynomial, and decays with a power equal to the number n of terms in the n -point function. Hence, if Bob neglects features whose distinguishability vanishes to order $2n$ in h_Π or h_Φ , the relevant observables are the products of n field operators

with momenta lower than $1/\sigma$.

For instance, for $n = 2$, only quadratic observables are relevant. Furthermore, since this statement holds in the neighbourhood of any gaussian state ρ , the manifold of all gaussian states is everywhere tangent to the relevant observables, and *orthogonal* to the irrelevant ones. Hence, it not only forms a good family of effective theories as it uniquely labels the equivalence classes, but it is also the most “economical” as it requires only the smallest changes in the state to accommodate a given set of new observable predictions.

Note that a threshold of $n = \infty$ also makes sense here and means that only the effective momentum cutoff $1/\sigma$ plays a role. This is equivalent to using the projective channel $\mathcal{E} = \text{Tr}_{|k| > 1/\sigma}$ which traces out modes with momenta larger than $1/\sigma$. This prescription defines a Hamiltonian version of *momentum-shell* renormalisation.

We know that adding a non-quadratic term to the effective hamiltonian typically yields divergent n -point functions, which can always be made finite by imposing an “unphysical” cutoff Λ on all (spatial) momentum integrals. However, in order for the state to stay on a given equivalence class when Λ is varied, one may need to work within a larger class of hamiltonians, i.e., to add new terms with new coupling constants to be varied as a function of Λ . If only finitely many new parameters are required we say that the theory is *renormalisable*, but this is not a fundamental requirement for the effective manifold to be useful.

From the “condensed matter” point of view, one has a physical cutoff Λ that is fixed and given by the lattice spacing. If we start with what we suppose to be the “true” state on the lattice, then for the purpose of making predictions that Bob can observe, the description of the state can be simplified by exploiting the freedom we have in choosing a representative of the equivalence class. Our analysis guarantees that, to first order in perturbation theory, we can simply set the coupling constant in front of irrelevant observables in the hamiltonian to zero (because $e^{-H+\epsilon A} \approx e^{-H} + \epsilon \Omega_{e^{-H}}(A)$). Since polynomials in the fields involving modes with $|k| > 1/\sigma$ are irrelevant, this is mathematically equivalent to lowering the cutoff to $\Lambda = 1/\sigma$. Hence the QFT-like RG flow in terms of regulator Λ is equal to the RG flow in terms of the noise parameter $1/\sigma$ in this example.

Scaling — One may be uncomfortable with viewing the cutoff Λ as an explicit parameter of the state. A change of cutoff from Λ_0 to Λ can be absorbed into a *rescaling* of space by a factor $s = \Lambda/\Lambda_0$. The requirement that the couplings run with Λ so that the n -point functions be constant now takes a slightly different form: the couplings must run with s in such a way that the n -point function depends on s via a scale transformation. This is the requirement expressed by the Callan-Symanzik equations.

Conclusion — We showed how, using any model of experimental limitations, one can use information metrics to determine the most experimentally relevant pa-

rameters in the neighbourhood of an initial hypothesis, and safely ignore the irrelevant ones. We validated this approach by showing that, when applied in the neighbourhood of a quantum gaussian field state and with a reasonable model of finite spatial and field-value experimental precision, it yields the usual renormalisation conditions in quantum field theory, but with the added bonus of justifying the special relevance of n -point functions for small n . This calculation also shows how the momentum cutoff arises naturally from a real-space coarse-graining approach. We believe that this work demystifies many aspects of the RG and opens the way for a deeper understanding of its powerful underlying principles. In partic-

ular, it clarifies the role played by information theory in the RG, which has been a long standing problem in this field.

Acknowledgments — Helpful discussions with numerous people are most gratefully acknowledged: a partial list includes Andrew Doherty, Jens Eisert, Steve Flammia, Jutho Haegeman, Gerard Milburn, Terry Rudolph, Tom Stace, Frank Verstraete, and Reinhard Werner. This work was supported by the ERC grants QFTCMPS and SIQS and by the cluster of excellence EXC 201 Quantum Engineering and Space-Time Research.

-
- [1] K. G. Wilson and J. B. Kogut, Phys. Rept. **12**, 75 (1974).
 - [2] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).
 - [3] M. E. Fisher, Rev. Modern Phys. **70**, 653 (1998).
 - [4] G. I. Barenblatt, *Scaling, self-similarity, and intermediate asymptotics: dimensional analysis and intermediate asymptotics*, Vol. 14 (Cambridge University Press, 1996).
 - [5] J. Preskill, J. Mod. Opt. **47**, 127 (2000), arXiv:quant-ph/9904022.
 - [6] L. P. Kadanoff, Ann. Phys. **100**, 359 (1976).
 - [7] L. P. Kadanoff, Physics **2**, 263 (1966).
 - [8] L. P. Kadanoff, Rev. Mod. Phys. **49**, 267 (1977).
 - [9] A. B. Zamolodchikov, JETP Lett. **43**, 730 (1986), Translated from Pis'ma Zh. Èksper. Teoret. Fiz. **43**, 565–567 (1986), no. 12, (Russian).
 - [10] B. B. Machta, R. Chachra, M. K. Transtrum, and J. P. Sethna, Science **342**, 604 (2013), arXiv:1303.6738.
 - [11] S. M. Apenko, Physica A **391**, 62 (2012), arXiv:0910.2097.
 - [12] D. C. Brody and A. Ritz, Nucl. Phys. B **522**, 588 (1998), arXiv:hep-th/9709175.
 - [13] H. Casini and M. Huerta, J. Phys. A **40**, 7031 (2007), arXiv:cond-mat/0610375.
 - [14] J. Gaiete and D. O'Connor, Phys. Rev. D **54**, 5163 (1996), arXiv:hep-th/9511090.
 - [15] F. Hiai and D. Petz, Communications in mathematical physics **143**, 99 (1991).
 - [16] D. Petz, Linear algebra and its applications **244**, 81 (1996).
 - [17] This is a generalisation of the *transpose channel* defined in Ref. [19].
 - [18] It is also the ratio of the original and coarse-grained *stiffness* of X_j , studied for classical models in Ref. [10].
 - [19] M. Ohya and D. Petz, *Quantum entropy and its use* (Springer Verlag, 2004).